LBNL-51905

Traditional Risk Factor Study (TRiFS): Preliminary Analysis of Risk Factor Prevalence and Population Attributable Fraction for Breast Cancer in Marin County, California

Erdmann CA[1], Satariano WA[2], Chen YQ[2]

[1]Lawrence Berkeley National Laboratory
[2]University of California at Berkeley

Report submitted to the
Marin County Department of Health and Human Services

Address correspondence to:
C. A. Erdmann, Ph.D., MPH
Mailstop 90-3058
Lawrence Berkeley National Laboratory
One Cyclotron Road
Berkeley, CA 94720
U.S.A.

May, 2003

This work was made possible by funds received from Agreement No. 00-91462 with the California Department of Health Services, Cancer Research Section administered through the Marin County Department of Health and Human Services. This project was also partially supported by the U.S. Department of Energy under No. DE-AC03-76SF00098.

DISCLAIMERS

ACKNOWLEDGEMENTS

Dr. Margaret Wrensch at the University of California at San Francisco provided data from the Marin County Breast Cancer Study of Adolescent Risk Factors (ARFS) and valuable comments on a draft of this report. Terri Chew provided documentation for the ARFS dataset. Colin Leary made major contributions to the initial research on the availability of datasets for this study. Breast cancer risk factor data for the State of California and data for county-level comparisons were provided by the 2001 California Health Interview Survey (CHIS). We especially thank Lee Habte at the Center for Health Policy Research at the University of California Los Angeles for assistance in navigating the CHIS data use process.

ABSTRACT

The objective of the Traditional Risk Factor Study (TRiFS) was to describe female breast cancer risk factor distributions in Marin County, California by using previously collected, individual-level data. The Marin County breast cancer risk factor distributions were compared with those of the other California counties and the State. Prevalence estimates for traditional breast cancer risk factors (e.g., age at menarche, family history, and age at first birth) were computed using data from the Marin County Breast Cancer Study of Adolescent Risk Factors (ARFS) and the California Health Interview Survey (CHIS). A reference set of relative risk values for the breast cancer risk factors of interest was assembled from published sources. Using the prevalence estimates along with these relative risk values, population attributable fractions were calculated for selected breast cancer risk factors and combinations of these factors. Approximately 84% of Marin County women were exposed to at least one of the following five breast cancer risk factors: earlier age at menarche, later age at 1st birth or nulliparity, family history, later age at menopause, and/or higher postmenopausal body mass index (BMI). The results suggest that 50% of Marin County's breast cancer cases would be avoided if the traditional breast cancer risk factors considered in these analyses were eliminated. Later age at first birth and nulliparity after age 30 alone appear to account for about one-third of breast cancer cases in Marin County.

INTRODUCTION

The San Francisco Bay Area (SFBA) reports some of the highest breast cancer incidence in the world. Within this region, breast cancer incidence and mortality are highest in Marin County, a small, urban county of 250,000 predominantly White, Non-Hispanic residents located immediately north of the city of San Francisco (Clarke et al., 2002). Averaged over the period 1995-1999, age-adjusted invasive breast cancer incidence per 100,000 for Non-Hispanic White women were 199 in Marin County, 155 in the rest of the SFBA, and 144 in the United States as a whole (Clarke et al., 2002). One possible explanation for the high incidence of breast cancer in Marin County is that known breast cancer risk factors are more prevalent in Marin than in areas of lower incidence. Therefore, we set out to estimate the proportion of breast cancer cases in Marin County that are attributable to traditional breast cancer risk factors. Using census block *group-level* data, a previous study found that the variation in breast cancer incidence among 25 California counties, including five San Francisco Bay Area counties (Alameda, Contra Costa, Marin, San Francisco, and San Mateo), was attributable to differences in levels of known risk factors for the years 1988-1992 (Prehn & West, 1998)[1]. The objective of the Traditional Risk Factor Study (TRiFS) was to describe female breast cancer risk factor distributions in Marin County using previously collected, *individual-level* data. This represents a cost-effective and efficient strategy to understand more about the risk of breast cancer among women in Marin County by using existing sources of data. Prevalence and population attributable fraction estimates for selected traditional breast cancer risk factors (see Table 1) were calculated and compared with those of other California counties and the State.

> ***Population Attributable Fraction*** is the "the proportion of disease cases over a specified time that would be prevented following elimination of the exposures, assuming the exposures are causal." - Rockhill et al. (1998)

Table 1. Traditional breast cancer risk factors considered in the TRiFS analyses.

Variable	Higher Risk Group Definition
Age at menarche	Age at menarche < 12 years
Age at 1st birth	Age at 1st birth ≥ 30 years (includes nulliparous women over age 30)
Parity	Nulliparous
Family history	Breast cancer in one or more 1st degree relatives*
Age at menopause	Age at menopause ≥ 55 years†
Postmenopausal BMI	Body mass index > 27 (among postmenopausal women only)
Education	Highest level of education at least a Bachelor's degree or equivalent
SES	Current socioeconomic status‡

*i.e., mother, sister, and/or daughter.
†Among postmenopausal women and women aged 55 years and older.
‡ SES was defined differently in the ARFS and CHIS datasets. These definition differences are described below in the DATASETS section.

[1]Cases were 30,289 women diagnosed with incident female invasive breast cancer in the years 1990-1994. This ecologic analysis used census block group as the unit of analysis and considered measures of age, parity, urban/rural status, percent of households living below poverty level, percent of persons 25 and older with a college education, median household income, median per capita income, percent of employed persons in a working class occupation, and marital status. Census block groups that had a high level of any risk factor had higher incidence rates, regardless of geographic location.

DATASETS

A key part of this project was to identify and summarize datasets having Marin-specific, individual-level breast cancer risk factor data and to select two of these datasets for further study. Summaries of these datasets were provided in an interim report (see Appendix A). The two datasets selected for further study were the Marin County Breast Cancer Study of Adolescent Risk Factors (ARFS) and the 2001 California Health Interview Survey (CHIS).

ARFS - ARFS provides the richest available Marin-specific dataset in terms of information about breast cancer risk factors. ARFS was a population-based case-control study of women aged 31-74 years residing in Marin County. Data were collected on generally recognized breast cancer risk factors and a variety of adolescent physical, psychological, and social factors. Participants included 336 cases and 321 controls that participated in either full in-person or abbreviated telephone interviews. Cases were female Marin County residents aged 31-74 diagnosed with primary breast cancer between July 1997 and June 1999. Contact information for the cases was obtained from the regional cancer registry operated by the Northern California Cancer Center (NCCC). Controls were female Marin County residents aged 31-74 years at the time of their participation in the study. Control women without breast cancer were ascertained through random digit dialing (RDD) and were frequency matched to cases by age at diagnosis (within five years) and ethnicity. The 300 cases and 305 controls that participated in full in-person interviews were included in the TRiFS analyses. In the ARFS dataset, the higher SES group consisted of women who self-reported their current SES as being Upper Middle or Upper class. Further detail regarding ARFS can be found in Wrensch et al. (2003).

CHIS - The dataset of the CHIS Adult Questionnaire was selected as a suitable complementary dataset as it includes breast cancer risk factor data at the county-level for counties other than Marin and the State. The CHIS was a large telephone survey that conducted health-related interviews with one randomly selected adult in each of 55,000 households sampled in California. Survey topics included health-related behaviors, health status and conditions, health insurance coverage, and access to health care services. Interviews were conducted in six languages: English, Spanish, Chinese (Mandarin and Cantonese dialects), Vietnamese, Korean, and Khmer (Cambodian). The CHIS sample was designed to provide population-based estimates for most California counties (including Marin County). In CHIS, the 58 California counties were arranged into 41 strata. Thirty-three of the 35 counties with a population of 100,000 or more formed their own strata. Monterey and Humboldt counties were combined with smaller adjoining counties (San Benito and Del Norte, respectively). The 23 counties with populations smaller than 100,000 were grouped into six strata. In addition, the cities of Long Beach, Pasadena, and Berkeley were treated as separate strata. For analysis of the CHIS dataset, the higher SES group was defined to consist of women who reported their annual household income to be greater than $30,000, which corresponds to the upper 2/3 cut-point for U.S. household incomes in 1999 (2000 U.S. Census). Age at menopause was not available in the CHIS dataset. To estimate postmenopausal BMI from the CHIS dataset, women aged 50 and older were assumed to be postmenopausal. Additional information regarding the CHIS, including a copy of the questionnaire and data dictionary, can be found at the CHIS website (www.chis.ucla.edu).

METHODS

Prevalence

Prevalence estimates for the selected traditional breast cancer risk factors were computed for the ARFS control group. Since ARFS was a case-control study, the proportion of cases in the total group of study participants is not representative of the proportion of cases in the Marin County female population. As breast cancer is a rare disease, the control group alone provides the best prevalence estimates for the source population when using case-control data. Source population prevalence estimates were tabulated for females aged 31-74 years for CHIS. This restricted age group corresponds to that of the ARFS participants. For CHIS, the population prevalence estimates were computed using the procedure PROC SURVEYMEANS in SAS (a statistical software package) to incorporate sample weights and account for the complex sample design. The syntax of this procedure is as follows:

```
proc surveymeans data = {SAS dataset};
  var {variables to be analyzed};
  class {variables to be analyzed as categorical variables};
  weight {sample weight variable};
  strata {variable that forms the strata in a stratified sample};
run;
```

The CHIS weight and strata variables used in these analyses were RAKEDW0 and STRATA_2, respectively. Since the ARFS control group was a population-based random sample, no sample weighting or design adjustment was necessary.

Population Attributable Fraction

To calculate population attributable fraction (PAF) estimates, we used the formula

$$PAF = \frac{p_e(RR-1)}{p_e(RR-1)+1} \qquad (1)$$

where p_e is the estimated proportion of the source population that is exposed to the factor of interest (i.e., prevalence), and RR is the relative risk estimate for the factor of interest ($RR \geq 1$); RR may be a risk ratio, a rate ratio, or an approximation of one of these two ratios, such as an odds ratio (Levin, 1971 in Miettinen, 1974; Cole and MacMahon, 1971; Rockhill et al., 1998). This formula allows for calculation of PAF without knowing the incidence rate in the population among the unexposed (Kleinbaum, 1982). An extension of formula 1 may be used to calculate the PAF for multiple-category exposures (Walter, 1976; Kleinbaum, 1982; Rockhill et al., 1998):

$$PAF = \frac{\sum_{j=0}^{k}(p_j)(RR_j-1)}{1+\sum_{j=0}^{k}(p_j)(RR_j-1)} = 1-\frac{1}{\sum_{j=0}^{k}p_j(RR_j)}. \qquad (2)$$

In formula 2, p_j is the estimated proportion of the source population in the j^{th} exposure category, RR_j is the relative risk estimate comparing the j^{th} exposure category with the unexposed group ($j = 0$). Defining each unique risk factor combination as an exposure category, formula 2 is used to calculate the PAF for combinations of risk factors. Formula 2 can be algebraically transformed such that rr_i is the relative risk estimate for the i^{th} individual and x is the total number of individuals. In other words, rr_i is the risk for each individual, based on their own set of risk factors, relative to an individual in the lowest risk, or unexposed, group ($j = 0$). Each individual in stratum j is assigned $rr_i = RR_j$, and the equation is re-written as

$$PAF = 1 - \frac{1}{\frac{1}{x}\sum_{i=1}^{x} rr_i} = 1 - \frac{x}{\sum_{i=1}^{x} rr_i}. \qquad (3)$$

Relative Risk Estimates

A reference set of relative risk values for the breast cancer risk factors of interest was assembled from published sources. Table 2 summarizes relative risk estimates used in these analyses. Not all sources consulted offered an estimate for all of the risk factors considered in these analyses. Relative risk estimates from the Gail Model were used whenever possible (Gail et al., 1989). The Gail Model is frequently used in the clinical setting to assess breast cancer risk for individual women (Willett et al, 2000). For other variables, a range of relative risk estimates was used. Nulliparous women were grouped with women whose first birth was at age 30 years or older since nulliparous women have approximately the same risk as women with a first birth around age 30 years (NCI, 2003).

Table 2. Reference set of relative risk estimates from published sources.

Risk Factor	RR	Source
Age at menarche <12 years	1.2	Gail et al., 1989
Age at 1st birth ≥30 years	1.9	Gail et al., 1989; NCI, 2003
Parity	1.9	Gail et al., 1989; NCI, 2003
Family history	2.6	Gail et al., 1989
Age at menopause ≥ 55 years	1.1-2.0	Kelsey, 1993
BMI (postmenopause) >27	1.1-1.4	Willett et al., 2000
Education ≥ B.A./B.S.	1.4-2.3	Horn-Ross et al., 2001
High SES	1.1-2.0	Kelsey, 1993

Using the source population prevalences along with the reference set of relative risk values, population attributable fractions were calculated for the selected breast cancer risk factors and combinations of these factors. SAS, Splus, and R statistical software packages were used for these analyses. PERL (a unix-based software tool) was used for some editing of the analysis data files.

RESULTS

Table 3 gives the source population prevalences for Marin County expressed as a percentage for various strata of the breast cancer risk factors considered in these analyses. Based on the ARFS dataset, we estimated that 84% of Marin County women had at least one of the following five breast cancer risk factors: earlier age at menarche, later age at 1st birth or nulliparity, family history, later age at menopause, and/or higher postmenopausal BMI. Approximately, 44% of Marin County women were exposed to two or more of the aforementioned five breast cancer risk factors. Prevalence estimates expressed as a percentage for the State of California and the 44 CHIS strata are given in Table 4.

Table 5 presents the PAF estimates for each of the seven traditional breast cancer risk factors considered in these analyses for the State of California and the 44 CHIS strata. Setting aside education and income, later age at first birth/nulliparity and family history were two single risk factors that each consistently accounted for 10 percent or more of breast cancer cases across the 44 CHIS strata, with PAF ranges of 10.3-36.5 and 11.0-28.9, respectively. Interestingly, there was little overlap across counties with respect to the five highest values for each risk factor, indicating that risk factor prevalence patterns vary across counties. That is, counties with a relatively high prevalence of one risk factor do not necessarily have a high prevalence of other risk factors. While there did not appear to be strong clustering of risk factors within certain counties, some moderate correlations of risk factor prevalences were noted. For example, the prevalence of later age at first birth was negatively correlated with earlier age at menarche ($r = -0.53$), indicating that early age at menarche may be associated with earlier age at first birth. Previous studies also have found early age at menarche to be related to early childbearing (Talashek et al., 2000; Udry, 1979).

Table 6 compares the PAF estimates obtained from ARFS and CHIS. While the Marin County PAF estimates derived from the ARFS and CHIS datasets did not substantially differ, there were some differences. Most notably, the Marin County ARFS-based PAF estimate for family history was greater than the CHIS-based estimate (24.2 versus 17.9). Differences in ARFS and CHIS estimates for family history may be due to participation bias issues. Since ARFS was specifically a breast cancer study, the controls may have been more willing to participate if they had a relative with breast cancer. Whereas, because CHIS was a general health survey, CHIS participants perhaps were less likely to base their decision to participate upon whether they were related to anyone with breast cancer. Also of note, the ARFS-based socioeconomic status (SES) PAF estimate for Marin County was smaller than that derived from the CHIS dataset (5.0-34.6 versus 7.7-45.3). Because SES was measured differently in the two datasets, the SES PAF estimates derived from ARFS and CHIS are not directly comparable. For the ARFS dataset, the higher SES group consisted of women who self-reported their current SES as being Upper Middle or Upper class. For the CHIS dataset, the higher SES group consisted of women who reported their annual household income to be greater than $30,000, which corresponds to the upper 2/3 cut-point for U.S. household incomes in 1999 (2000 U.S. Census).

Marin County PAF estimates were higher than the state average for later age at first birth/nulliparity, family history, education, and income. Later age at first birth and

nulliparity after age 30 alone appears to account for about one-third of breast cancer cases in Marin County (Table 6). Considered in combination, traditional breast cancer risk factors, excluding income and education, appear to account for about half of the breast cancer cases in Marin County (Table 7).

Table 3. Marin County source population prevalence estimates from ARFS and CHIS datasets for various strata of selected breast cancer risk factors.

Risk Factor	% ARFS Controls aged 31-74 years	% CHIS females aged 31-74 years
Age at menarche		
≥14	20	26
12-13	58	59
<12	22	15
Age at 1st birth		
<20	5	9
20-29	62	56
≥30	33	35
Age at 1st birth		
<30	49	42
≥30 (includes nulliparous)	51	58
Parity		
Parous	73	66
Nulliparous	27	34
Family history		
No	80	86
Yes	20	14
Age at menopause		
<45	22	N.A.*
45-54	64	N.A.*
≥55	14	N.A.*
BMI (postmenopausal)†		
≤27	60	72
>27	40	28
Education		
< B.A/B.S.	40	38
≥ B.A./B.S.	60	62
SES‡		
Lower	47	17
Higher	53	83

*Age at menopause was not available in the CHIS dataset.

†Since menopausal status was not available in the CHIS dataset, women aged 50 and older were assumed to be postmenopausal to estimate postmenopausal BMI from the CHIS dataset.

‡For the ARFS dataset the higher SES group consisted of women who self-reported their current SES as being Upper Middle or Upper class. For the CHIS dataset, the higher SES group consisted of women who reported their annual household income to be greater than $30,000, which corresponds to the upper 2/3 cut-point for U.S. household incomes in 1999 (2000 U.S. Census).

Table 4. Percentage of the source population with selected breast cancer risk factors based on the CHIS dataset for the State of California and the 44 CHIS strata. The shaded cells contain the five highest values for each variable.

	State/County/Strata	Age at menarche	Age at 1st birth	Parity	Family Hx	Postmenopausal BMI	Education	Income
0	CALIFORNIA STATE*	19	31	17	12	41	32	67
1.1	Long Beach	23	30	21	11	32	29	57
1.2	Pasadena	21	49	31	19	32	50	71
1.3	Los Angeles Balance	21	32	18	10	43	29	60
2	San Diego	16	31	16	11	35	35	69
3	Orange	16	29	17	11	39	37	75
4	Santa Clara	16	38	18	11	37	45	77
5	San Bernadino	24	19	9	13	43	24	62
6	Riverside	20	19	12	12	44	20	63
7.1	Berkeley	14	64	46	17	15	81	84
7.2	Alameda Balance	19	40	20	15	38	42	74
8	Sacramento	19	27	16	14	42	28	72
9	Contra Costa	19	37	20	16	37	43	83
10	Fresno	20	21	12	14	58	22	54
11	San Francisco	17	57	38	12	31	52	66
12	Ventura	20	30	15	15	42	30	75
13	San Mateo	17	43	23	17	36	43	80
14	Kern	21	16	9	11	49	13	53
15	San Joaquin	19	19	13	12	53	23	62
16	Sonoma	19	37	22	13	35	41	76
17	Stanislaus	17	18	8	11	51	18	61
18	Santa Barbara	24	26	13	9	34	30	66
19	Solano	22	26	15	12	47	30	75
20	Tulare	25	18	7	11	57	14	49
21	Santa Cruz	16	36	17	17	29	41	74
22	**MARIN**	15	58	34	14	28	62	83
23	San Luis Obispo	19	35	20	24	42	30	68
24	Placer	19	31	14	12	36	35	82
25	Merced	23	17	9	13	57	12	54
26	Butte	23	27	15	19	33	29	55
27	Shasta	20	18	10	16	44	19	59
28	28. Yolo	17	32	18	16	37	44	70
29	El Dorado	18	27	14	21	33	29	75
30	Imperial	17	17	9	8	50	13	41
31	Napa	14	32	18	18	38	33	79
32	Kings	24	13	7	9	54	14	50
33	Madera	22	18	9	12	40	17	57
34	Monterey San Benito	21	22	14	9	40	23	64
35	Del Norte Humboldt	22	27	16	13	47	32	59
36	Lassen, etc.	20	20	12	17	43	19	53
37	Lake Mendocino	19	23	13	20	45	19	51
38	Colusa Glenn Tehama	19	15	9	17	47	12	49
39	Sutter Yuba	23	18	10	11	54	14	55
40	Nevada Plumas Sierra	18	30	16	23	28	33	76
41	Alpine, etc.	16	23	12	25	34	20	64

*including Marin County

Table 5. Population attributable fraction estimates (expressed as a percentage) based on the CHIS dataset for the State of California and the 44 CHIS strata. The shaded cells contain the five highest values for each variable.

	State/County/Strata	Age at menarche	Age at 1st birth	Parity	Family Hx	Postmenopausal BMI†	Education†	Income†
0	CALIFORNIA STATE*	3.7	21.5	13.3	16.4	3.9-14.0	11.3-29.2	6.3-40.0
1.1	Long Beach	4.3	21.3	15.6	15.5	3.1-11.4	10.5-27.5	5.4-36.2
1.2	Pasadena	4	30.6	21.8	23.4	3.1-11.2	16.7-39.5	6.6-41.4
1.3	Los Angeles Balance	4	22.1	14.2	13.8	4.1-14.6	10.3-27.1	5.6-37.4
2	San Diego	3.2	21.9	12.6	15.2	3.3-12.2	12.1-31	6.5-41
3	Orange	3.1	20.9	13	15	3.8-13.5	12.8-32.4	7-42.9
4	Santa Clara	3	25.6	14.1	14.5	3.6-12.9	15.3-36.9	7.2-43.6
5	San Bernadino	4.5	14.5	7.7	17.7	4.1-14.7	8.8-24	5.9-38.4
6	Riverside	3.8	14.5	9.5	15.9	4.3-15.1	7.5-20.8	5.9-38.7
7.1	Berkeley	2.7	36.5	29.2	21.5	1.5-5.8	24.4-51.2	7.8-45.7
7.2	Alameda Balance	3.6	26.6	15.5	19	3.7-13.2	14.3-35.1	6.9-42.4
8	Sacramento	3.7	19.8	12.5	18.8	4-14.4	10.2-27	6.7-41.8
9	Contra Costa	3.7	24.8	15.3	20.5	3.6-12.8	14.7-35.9	7.7-45.5
10	Fresno	3.9	15.7	9.8	18.6	5.4-18.7	8.2-22.5	5.2-35.2
11	San Francisco	3.2	33.9	25.4	15.9	3-11.2	17.2-40.2	6.2-39.9
12	Ventura	3.9	21.4	11.6	19.3	4-14.3	10.7-28.1	7-42.9
13	San Mateo	3.3	28	17.2	21.8	3.5-12.5	14.7-35.8	7.4-44.6
14	Kern	4	12.6	7.7	15.4	4.7-16.5	5-14.5	5-34.6
15	San Joaquin	3.7	14.6	10.6	15.7	5.1-17.6	8.5-23.1	5.8-38.1
16	Sonoma	3.7	25.1	16.2	17.3	3.4-12.3	14.1-34.8	7-43
17	Stanislaus	3.2	14.2	6.9	15.1	4.9-17	6.6-18.6	5.7-37.8
18	Santa Barbara	4.5	19.2	10.2	12.3	3.3-11.9	10.6-27.8	6.2-39.9
19	Solano	4.2	19.2	12	15.9	4.5-15.9	10.6-27.7	7-43
20	Tulare	4.7	13.7	5.7	15	5.4-18.5	5.2-15.1	4.7-32.8
21	Santa Cruz	3.1	24.7	12.9	21.8	2.8-10.3	14-34.5	6.9-42.6
22	**MARIN**	3	34.1	23.6	17.9	2.7-10.1	19.8-44.6	7.7-45.3
23	San Luis Obispo	3.6	23.8	15	27.6	4.1-14.5	10.8-28.2	6.4-40.6
24	Placer	3.7	21.8	11.4	15.9	3.5-12.7	12.3-31.4	7.6-45.1
25	Merced	4.5	13.1	7.8	17	5.4-18.6	4.7-13.9	5.2-35.2
26	Butte	4.4	19.4	11.8	23.2	3.2-11.8	10.5-27.6	5.2-35.6
27	Shasta	3.9	14.2	8.6	20.5	4.2-14.9	7-19.8	5.6-37.1
28	28. Yolo	3.2	22.2	14	20.5	3.6-12.9	14.8-36.2	6.5-41.2
29	El Dorado	3.4	19.4	11.1	25.6	3.2-11.6	10.6-27.7	7-42.9
30	Imperial	3.3	13.1	7.5	11	4.7-16.6	4.8-14	3.9-29.1
31	Napa	2.7	22.6	13.7	22.3	3.6-13.1	11.6-29.8	7.3-44.1
32	Kings	4.6	10.3	5.6	12.6	5.1-17.7	5.3-15.3	4.7-33.2
33	Madera	4.3	14.2	7.5	16.3	3.8-13.8	6.4-18.1	5.4-36.3
34	Monterey San Benito	4.1	16.5	11.4	12.6	3.8-13.7	8.4-23	6-39.1
35	Del Norte Humboldt	4.1	19.7	12.4	17.1	4.5-15.9	11.2-29.1	5.6-37.3
36	Lassen, etc.	3.9	15	9.5	21.8	4.1-14.7	7.2-20.1	5-34.6
37	Lake Mendocino	3.6	17	10.3	24.7	4.3-15.3	7.2-20.2	4.9-33.8
38	Colusa Glenn Tehama	3.6	12.1	7.3	20.9	4.5-15.8	4.7-13.8	4.7-33
39	Sutter Yuba	4.4	14.1	8.4	15.2	5.1-17.6	5.2-15	5.2-35.6
40	Nevada Plumas Sierra	3.5	21.4	12.3	27.1	2.7-10	11.8-30.3	7-43.1
41	Alpine, etc.	3.1	17.1	10	28.9	3.3-12	7.6-21	6-38.9

*including Marin County

†The ranges that appear in these columns are not confidence intervals. Rather they reflect the range of estimates obtained using the extremes of the RR estimates used for these calculations (see Table 2).

Table 6. Population attributable fraction estimates (expressed as a percentage), comparing estimates from the ARFS and CHIS datasets.

Source Population	Age at menarche	Age at 1st birth	Parity	Family History	Age at Menopause†	Postmenopausal BMI†	Education†	SES†‡
ARFS	4.2	31.3	19.3	24.2	1.4-12.0	3.9-13.9	19.3-43.7	5.0-34.6
CHIS Marin	3.0	34.1	23.6	17.9	N.A.	2.7-10.1	19.8-44.6	7.7-45.3
CHIS State*	3.7	21.4	13.2	16.4	N.A.	3.9-14.0	11.1-29.0	6.2-40.0

*excluding Marin County

†The ranges that appear in these columns are not confidence intervals. Rather they reflect the range of estimates obtained using the extremes of the RR estimates used for these calculations (see Table 2).

‡SES was measured differently in ARFS and CHIS, and therefore the PAF estimates made using ARFS data are not directly comparable to those made using the CHIS data.

Table 7. Population attributable fraction estimates (expressed as a percentage) for combinations of breast cancer risk factors risk factors calculated using the ARFS† dataset.

Risk Factor Combination	RR	PAF*
Age at menarche	1.2	51.3-57.6
Age at 1st birth	1.9	
Family history	2.6	
Age at menopause	1.1-2.0	
Postmenopausal BMI	1.1-1.4	
Age at 1st birth	1.9	32.8-36.9
Postmenopausal BMI	1.1-1.4	

*The ranges that appear in these columns are not confidence intervals. Rather they reflect the range of estimates obtained using the extremes of the RR estimates used for these calculations (see Table 2).

†Record-level data were not available from the CHIS dataset to calculate PAF estimates for risk factor combinations at the time of these analyses.

SUMMARY

Using previously collected individual-level data, this report provides prevalence and population attributable fraction estimates for traditional breast cancer risk factors in Marin County, other California counties, and the State of California. The risk factors included in these analyses were selected based upon their recognition as established breast cancer risk factors in review articles at the project's initiation and data availability. The results from this study suggest that approximately 50% of Marin County's breast cancer cases would be avoided if the breast cancer risk factors considered in these analyses were eliminated from the Marin population. Since most of these risk factors are not readily modifiable, a better understanding of the biologic mechanisms of these factors may lead to insights into modifiable components of these factors. Researchers in the U.S. and elsewhere continue to develop further understanding of the role of modifiable factors such as hormone replacement therapy use, alcohol consumption, and other environmental factors.

REFERENCES

Clarke CA, Glaser SL, West DW, Ereman R, Erdmann C, Barlow J, Wrensch M. Breast Cancer incidence and mortality trends in an affluent population: Marin County, California, USA, 1990-1999. Breast Cancer Research 2002,4:R13.

Cole P, MacMahon. Attributable risk percent in case-control studies. Brit J Prev Soc Med 1971;25:242-244.

Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Mulvihill JJ. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. Journal of the National Cancer Institute 1989;81(24):1879-1886.

Horn-Ross PL, John EM, Lee M, Stewart SL, Koo J, Sakoda LC, Shiau AC, Goldstein J, Davis P, Perez-Stable EJ. Phytoestrogen consumption and breast cancer risk in a multiethnic population: The Bay Area Breast Cancer Study. American Journal of Epidemiology 2001;154(5):434-441.

Kelsey JL. Breast cancer epidemiology: Summary and future directions. Epidemiologic Reviews 1993;15(1):256-263.

Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic Research: Principles and Quantitative Methods. New York: Van Nostrand Reinhold, 1982.

Levin ML. The occurrence of lung cancer in man. Acta Un Intern Cancer 1953;9:531-541.

Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait or intervention. American Journal of Epidemiology 1974;99(5):325-332.

National Cancer Institute (NCI). Summary Report: Early Reproductive Events and Breast Cancer Workshop. http://www.nci.nih.gov/cancerinfo/ere-workshop-report (last accessed on 5/6/2003).

Prehn AW, West DW. Evaluating local differences in breast cancer incidence rates: A census-based methodology (United States). Cancer Causes and Control 1998;9:511-517.

Rockhill B, Weinberg CR, Newman B. Population attributable fraction estimation for established breast cancer risk factors: Considering the issues of high prevalence and unmodifiability. American Journal of Epidemiology 1998;147;9:826-833.

Talashek ML, Montgomery AC, Moran C, Paskiewicz L, Jiang Y. Menarche, sexual practices, and pregnancy: model testing. Clin Excell Nurse Pract 2000;4(2):98-107.

Udry JR. Age at menarche, at first intercourse, and at first pregnancy. J Biosoc Sci 1979;11(4):433-41.

Willett WC, Rockhill B, Hankinson SE, Hunter DJ, Colditz GA. Epidemiology and assessing and managing risk. In: Harris JR, Lippman ME, Morrow M, Osborne CK. Diseases of the Breast. New York: Lippincott Williams & Wilkins, 2000.

Wrensch M, Chew T, Farren G, Barlow J, Belli F, Clarke C, Erdmann CA, Lee M, Moghadassi M, Peskin-Mentzer R, Quesenberry Jr. CP, Souders-Mason V, Spence L, Suzuki M, Gould M. Risk factors for breast cancer in a population with high incidence rates. Breast Cancer Research 2003;5(4):R88-R102.

APPENDICES

APPENDIX A - Interim report

APPENDIX B - Slides from presentation given at Marin County Town Meeting on 5/29/03

APPENDIX A - Interim report

LBNL-51905

Traditional Risk Factor Study (TRiFS)
Dataset Research & Preliminary Data Analysis Plan

---Interim Report---

Christine A. Erdmann, Ph.D.
Co-Principal Investigator
Lawrence Berkeley National Laboratory

William Satariano, Ph.D.
Co-Principal Investigator
University of California at Berkeley

Ying Qing Chen, Ph.D.
Co-Investigator
University of California at Berkeley

Colin Leary
Research Assistant
Lawrence Berkeley National Laboratory

December 2002

This work was made possible by funds received from Agreement No. 00-91462 with the California Department of Health Services, Cancer Research Section administered through the Marin County Department of Health and Human Services. This project was also partially supported by the U.S. Department of Energy under No. DE-AC03-76SF00098.

Appendix A

DISCLAIMER

This document may contain research results which are experimental in nature. Neither the United States Government, nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not constitute or imply an endorsement or recommendation by the United States Government or any agency thereof, or by The Regents of the University of California. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or by The Regents of the University of California and shall not be used for advertising or product endorsement purposes.

Mention of trade name, proprietary product or specific equipment does not constitute a guaranty or warranty by the Department of Health Services, nor does it imply approval to the exclusion of other products. The views expressed herein represent those of the authors and do not necessarily represent the position of the State of California, Department of Health Services.

INTRODUCTION

The objective of the Traditional Risk Factor Study (TRiFS) is to describe female breast cancer incidence distributions in Marin County using previously collected data. The Marin County breast cancer risk distributions will be compared with those of California counties and the State. A key part of this project has been to identify datasets having Marin-specific, individual-level breast cancer risk factor data. The datasets identified to date are summarized in the Appendix. For each dataset, the following summary information has been assembled from the dataset documentation and reports: Dataset Name, Dataset Source, Procedure for Obtaining Dataset, Study Description/Abstract, Years of Data Collection, Study Design, Description of Participants, Data Collection Procedures, Response Rates, and Bibliography of Associated Reports. As part of this project, two of the identified datasets were selected for further study. The remainder of this interim report discusses the selected datasets and the preliminary plan for analyzing them.

SELECTION OF TWO DATASETS FOR FURTHER STUDY

The richest Marin-specific dataset, in terms of breast cancer risk factor data, is that of the Marin County Breast Cancer Study of Adolescent Risk Factors (ARFS). While the Buck Center Health and Functioning in Marin Study and the Marin County Health Survey both had a larger number of participants, the information on breast cancer risk factors in these datasets is limited. Furthermore, the Buck dataset study population focused exclusively on women aged 55 years and older (i.e., nearly all postmenopausal). In contrast, the ARFS dataset included both pre- and postmenopausal women, and the ARFS study criteria allowed for inclusion of women aged 21-75 years. In addition, the principal investigators of the ARFS study have graciously provided access to these data. Therefore, the ARFS dataset was the first to be selected for further study.

Selection of the second dataset was made in consultation with staff of the Marin County Breast Cancer Research Program and other members of the Marin Breast Cancer Research Collaborative. Table 1 summarizes key characteristics of the remaining datasets (i.e., excludes ARFS) so as to facilitate the decision-making process. Since part of the objective of TRiFS is to compare Marin breast cancer risk factor distributions with those of other California counties and the state and the ARFS dataset contains only Marin-specific data, one desirable feature of the second dataset is that it includes statewide and county-level data for counties other than Marin. Other key dataset characteristics considered in Table 1 are: number of Marin County participants (particularly women), time period covered, and which selected risk factors were available. The selected risk factor list included: parity, age at first full-term birth, BMI, age at menarche, age at menopause, family history, education, and socioeconomic status (SES)[1]. Through a process of elimination of the other remaining datasets, the California Health Interview Survey (CHIS) was selected as the second dataset. Table 2 summarizes the key reason(s) for elimination of each of the other datasets.

Two Datasets Selected for Further Study

1. Marin County Breast Cancer Study of Adolescent Risk Factors (ARFS)
2. California Health Interview Survey (CHIS)

[1] The datasets are summarized by the availability of a broader range of factors in the Appendix.

Table 1. Key characteristics of datasets to facilitate selection for further study.

Dataset[2]	Key Characteristics
ARFS	-Marin-specific data only -657 Marin women aged 31-74 years -Time period: 1999-2001 -Selected BC risk factors: parity, age at first full-term birth, BMI, age at menarche, age at menopause, family history, education, and SES
BRFS	-Statewide -184 Marin residents for 1997-2001 (men & women) -Time period: 1984-present -Selected BC risk factors: BMI, education, income, SES
Buck	-Marin-specific data only -1,162 Marin women aged 55 years and older -Time period: 1989-1991 -Selected BC risk factors: BMI, age at menopause, SES, income, education
CASH	-San Francisco Bay Area SEER counties only[3] -865 Bay Area women aged 20-54 years -Time period: 1980-1983 - Selected BC risk factors: parity, age at 1st full-term pregnancy, age at menarche, age at menopause, education, BMI, family history
CHIS	-Statewide -755 Marin residents aged 18 years and older (~350 women) -Time period: 2000-2001 -Selected BC risk factors: parity, age at first full-term birth, age at menarche, family history, education, and SES
CTS	-Statewide -15,660 Bay Area women (probably ~1,000 Marin women) -Time period: 1995/96, 1997, 2000 -Selected BC risk factors: parity, age at first full-term birth, BMI, age at menarche, age at menopause, family history
CWHS	-Statewide -186 Marin women aged 18 years and older for 1997-2001 -Time period: 1997-present -Selected BC risk factors: parity, age at 1st full-term pregnancy, education, income, SES, BMI
FRBC	-San Francisco Bay Area only (for California registry) -Minimal Marin-specific data available, only high risk women -Time period: 1996-present -Selected BC risk factors: parity, age at 1st full-term pregnancy, age at menarche, age at menopause, education, BMI
MCHS	-Marin-specific data only -4,821 Marin residents aged 18 years and older were surveyed -Time period: 2001 -Selected BC risk factors: parity, age at first full-term birth, BMI, family history, SES

[2] BRFS = Behavioral Risk Factor Study; Buck = Buck Center study of "Health and Functioning in Marin County"; CASH = Cancer and Steroid Hormone Study; CHIS = California Health Interview Study; CTS = California Teachers Study; CWHS = California Women's Health Study; FRBC = Family Registry for Breast Cancer; MCHS = Marin County Health Survey

[3] Alameda, Contra Costa, Marin, San Francisco, and San Mateo

Table 2. Datasets that were not selected for further study and the reason(s) for elimination.

Dataset[4]	Reason(s) for Elimination
BRFS	-Limited number of BC risk factors available -Limited number of Marin participants
Buck	-Marin-specific data only -Participants limited to age 55 years and older
CASH	-Limited to San Francisco Bay Area SEER counties only[5] -Participants limited to under age 55 years. -Data is a decade old
CTS	-Limited occupational group (teachers only) -Not a public use data set
CWHS	-Limited number of Marin participants
FRBC	-Limited number of Marin participants -Limited to only high risk women -Limited to San Francisco Bay Area only
MCHS	-Marin-specific data only

PRELIMINARY DATA ANALYSIS PLAN

ARFS - Marin County prevalence and population attributable fraction (PAF) estimates will be calculated for the following standard breast cancer risk factors: parity, age at first full-term birth, BMI, age at menarche, age at menopause, family history, education, and socioeconomic status.

CHIS – State and county-level prevalences and population attributable fraction (PAF) estimates will be calculated for the following standard breast cancer risk factors: parity, age at first full-term birth, age at menarche, family history, education, and socioeconomic status. Prevalences and PAF estimates may be calculated for additional variables (e.g., hormone replacement therapy use, oral contraceptive use, and alcohol use). The Dataset Summary Table in the Appendix summarizes the breast cancer-related variables available from the CHIS dataset.

Population Attributable Fraction (PAF) Estimates - A reference set of relative risk values for the standard breast cancer risk factors will be selected from the literature for calculating PAF estimates. Formulae recommended by Rockhill et al. (1998)[6] will be used. In addition to calculating PAF estimates for individual risk factors, we are developing methods to calculate PAF estimates for combinations of risk factors. Both SAS and R, a freely available form of the Splus language, will be used for all analyses.

---

[4] BRFS = Behavioral Risk Factor Study; Buck = Buck Center study of "Health and Functioning in Marin County"; CASH = Cancer and Steroid Hormone Study; CTS = California Teachers Study; CWHS = California Women's Health Study; FRBC = Family Registry for Breast Cancer; MCHS = Marin County Health Survey

[5] Alameda, Contra Costa, Marin, San Francisco, and San Mateo

[6] Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. American Journal of Public Health 1998;88(1):15-19.

APPENDIX

Traditional Risk Factor Study Dataset Descriptions

Appendix Table of Contents

Dataset Name:

Adolescent Risk Factor Study (ARFS)

Dataset Source:

Margaret Wrensch, Ph.D., Professor of Epidemiology, UCSF Department of Epidemiology and Biostatistics

Procedure for Obtaining Dataset:

This dataset is not a public use dataset.

Study Description/Abstract:

The ARFS is a population-based interview study of women with breast cancer and frequency age-matched control women without breast cancer residing in Marin County. It examined generally recognized breast cancer risk factors, as well as, a variety of adolescent physical development measurements, adolescent stress and social connection indicators, socioeconomic factors, passive and active smoking histories, alcohol usage, and childhood and adolescent residency. In one report based on this dataset, women with and without breast cancer are compared in terms of generally recognized breast cancer risk factors, childhood and adolescent socioeconomic factors, and several measures of health care utilization. In another, particular adolescent social experiences and exposures are analyzed.

Years of Data Collection:

1999-2001

Study Design:

Design:
Population-based case-control

Case ascertainment:
Eligible women included any resident of Marin County with diagnosis of primary breast cancer from July 1, 1997 through June 30, 1999 if under age 50 and through March 30, 1999 if age 50-75 at diagnosis. Names and addresses of cases were obtained through the Northern California Cancer Center Registry (a SEER registry).

Control ascertainment:
Control women (without breast cancer) were ascertained through random digit dialing and were frequency matched to cases' ages at diagnosis (within 5 years) and ethnicities. Seed numbers used for dialing included the area code plus first 6 telephone digits of all cancer cases in Marin County, 1998. Random digit dialing and preliminary eligibility screenings were conducted with 3,945 potential controls.

Follow-up:
A follow-up questionnaire was administered in order to collect buccal cell samples and obtain a complete Marin residential, educational, and workplace history. (Though there was some attrition for the follow-up, none of the variables in the follow-up questionnaire are pertinent to the aims of the TRiFS project).

Description of Participants:

Female participants by age and ethnicity:

Race/Ethnicity	25-34 y/o	35-44 y/o	45-54 y/o	55-64 y/o	65+ y/o	Total
White	3	69	209	201	104	**586**
Black	0	0	0	0	0	**0**
Hispanic	0	1	2	1	0	**4**
Asian	0	1	2	3	0	**6**
Mixed/Other	0	2	4	3	1	**10**
Sample Total	**3**	**73**	**217**	**208**	**105**	**606**

Marin Residency Breakdown:
All Participants were residents of Marin County at the time of diagnosis (cases) or recruitment (controls).

Data Collection Procedures:

In-person interviews:
The interviews were approximately two hours in length and conducted in-person at a place of each participant's choosing.

Response Rates:

Cases:
401 cases met the eligibility criteria. Of these, 301 (75%) completed the full interview, and 36 (9%) completed abbreviated telephone interviews.

Controls:
Of 347 women identified as eligible, 328 were contacted. Of the 328, 305 (93%) completed full interviews, and 16 (5%) completed the abbreviated telephone interview.

Abbreviated interview:
Some women who did not wish to participate in the full interview agreed to a brief, alternative telephone interview that included years in current residence, years in Marin County, place of birth, age at menarche, residence at age 11, whether family owned or rented residence at age 11, age first started working, highest grade or degree self and father achieved, selected major life events before age 21, ethnic background, family's socioeconomic status prior to age 21, highest education achieved, relative and actual body size, abbreviated histories of pregnancy, smoking, drinking, oral contraceptive and hormone use, and first degree relatives with breast cancer.

Bibliography of Associated Reports:

Wrensch M, Chew T, Farren G, Barlow J, Belli F, Clarke C, Erdmann CA, Lee M, Peskin-Mentzer R, Quesenberry CP, Souders-Mason V, Moghadassi M, Spence L, Suzuki M, Gould M. Breast cancer risk factors in a high risk population. 2002. (submitted).

Dataset Name:

Behavioral Risk Factor Survey (BRFS)

Dataset Source:

The survey instrument, technical documentation, and datasets may be obtained from the Survey Research Group at srg@ccr.ca.gov.

Procedure for Obtaining Dataset:

Data requests are to be made of:
Dr. Bonnie Davis
Chief, SRG
Cancer Surveillance System
Ph: 916-779-0331
bonnie@ccr.ca.gov

Study Description/Abstract:

The objective of the BRFS is to collect uniform, state-specific data on preventive health practices and risk behaviors that are linked to chronic diseases, injuries, and preventable infectious diseases in the adult population. Factors assessed by the BRFS include tobacco use, health care coverage, HIV/AIDS, physical activity, and diet.

Years of Data Collection:

Annually since 1984.

Study Design:

Data are collected from a random sample of non-institutionalized adults through a telephone survey. California used the Waksberg method of participant selection from 1984-1993 and used a disproportionate stratified sample (DSS) design from 1994-present.

Description of Participants:

Female participants by age and ethnicity:

Sex	Race/Ethnicity	18-24 y/o	25-34 y/o	35-44 y/o	45-54 y/o	55-64 y/o	65+ y/o	Total
Female	**White**	1528	3520	4235	3202	2414	4343	19242
	Black	244	514	453	282	197	210	1900
	Hispanic	1008	2077	1541	722	369	422	6139
	Asian/Other	305	613	527	297	169	142	2053
	Female subtotal	3085	6724	6735	4503	3149	5117	**29334**
	Sample Total	5723	12436	12198	8216	5561	8144	**52278**

Marin Residency Breakdown:

2001: n=41 Marin residents
2000: n=38 Marin residents
1999: n=36 Marin residents
1998: n=38 Marin residents
1997: n=31 Marin residents

Please note that the Marin residency figures are for both males and females. Marin figures are based on what is available in the published reports only. Further refinement is possible only after obtaining the datasets.

Data Collection Procedures:

Telephone interviews:
Starting in 1986, the survey was administered through computer-assisted telephone interviewing (CATI). Prior to 1986, data were recorded using pencil and paper. Following specifications provided by CDC, state health personnel or contractors conduct interviews in English or Spanish. The core portion of the questionnaire lasts an average of 10 minutes. Interview time for modules and state-added questions varies but generally extend an interview period by an additional 5 to 10 minutes. Telephone interviewing is conducted during a two-week period each month, 7 days per week, during both day and evening hours. Calls are rotated over days of the week and times of the day.

Response rates:

The CDC began response-rate calculations in 1987. The percentages of eligible households contacted which provided a completed interview follow:

Year	RR		Year	RR		Year	RR
1987	77%		1992	84%		1997	65%
1988	80%		1993	79%		1998	75%
1989	83%		1994	76%		1999	82%
1990	82%		1995	70%		2000	66%
1991	81%		1996	66%		2001	[unreported]

Bibliography of Associated Reports:

Remington PL, Smith MY, Williamson DF, Anda RF, Gentry EM, Hogelin GC. Design, characteristics, and usefulness of State-based Behavioral Risk Factor Surveillance: 1981-87. Public Health Reports 1988;103(4): 366-375.

Dataset Name:

Buck Baseline: "Health and Functioning in Marin County"

Dataset Source:

Buck Institute for Age Research
8001 Redwood Blvd.
Novato, CA 94945
Tel: 415-209-2000
Fax: 415-899-1810
Website: www.buckinstitute.org

Procedure for Obtaining Dataset:

This dataset is not a public use dataset.

Study Description/Abstract:

The Buck Center study was an effort to develop preliminary data concerning the onset and progression of common age-associated diseases and disorders (e.g., arthritis, sensory deficits, motor-sensory disorders), with the goal of assessing functional limitations associated with each health condition. It also obtained prevalence estimates of other common chronic and concurrent health conditions.

Years of Data Collection:

1989-1991

Study Design:

An age-stratified sample of non-institutionalized Marin County residents aged 55 and older was ascertained through random-digit-dialing of Marin County residential telephone numbers. The target sample was divided into four equal age groups: 55-64, 65-74, 75-84, and 85+,with approximately 500 Participants per age group. Older populations were oversampled using a list of Medicare-eligible residents.

Description of Participants:

Female participants by age:

	55-64 y/o	65-74 y/o	75-84 y/o	85+ y/o	Total
No. Participants	230	306	287	339	**1,162**

Marin Residency Breakdown:

All Participants were residents of Marin County.

Data Collection Procedures:

Agreeable participants arranged in-home interviews conducted by a team of trained volunteers.

Response Rates:

	55-64 y/o	65-74 y/o	75-84 y/o	85+ y/o	Total
% Participants	78.2	71.9	64.1	64.1	**69.4**

These response rates were calculated using the following equation:

$$\frac{\text{Completed interviews x 100}}{\text{Refusals + Completes}} = \text{Response rate}$$

Bibliography of Associated Reports:

Dataset Name:

Cancer and Steroid Hormone Study (CASH)

Dataset Source:

Centers for Disease Control and Prevention

Procedure for Obtaining Dataset:

Contract Polly Marchbanks (pam2@cdc.gov) to request a "mini-proposal" application form. Applications are reviewed by small committee.

Study Description/Abstract:

The Cancer and Steroid Hormone Study (CASH) was a multicenter, population-based, case-control study of oral contraceptive use in relation to breast, endometrial, and ovarian cancers diagnosed during the 1980-1982 time period. CASH study collected data in eight SEER regional registries: the metropolitan areas of Atlanta, Detroit, San Francisco, and Seattle; the states of Connecticut, Iowa, and New Mexico; and the four urban counties of Utah. These centers conducted case ascertainment, interviews, and retrieval of histologic slides of specimens, and medical records data. Risk factors tallied included education, parity, abortion history, breast-feeding, hormone usage, age at menarche/menopause, hysterectomy history, BMI, alcohol/tobacco usage, family history, and screening history.

Years of Data Collection:

December, 1980 through Apri,l 1983.

Study Design:

Design:
Population-based case-control

Cases:
Cases were women aged 20-54 years who resided in the eight study locations and who were newly diagnosed with primary breast, ovarian, or endometrial cancer between December 1, 1980 and December 31, 1982.

Controls:
Controls were women aged 20-54 years and selected through random digit dialing of households with telephones in the same eight geographic locations. The entire pool of controls was frequency-matched by the geographic location and 5-year age distribution of breast cancer cases.

Description of Participants:

Female participants by age:
Not available in the reviewed published reports. The following table describes selected cancer frequencies in the San Francisco Bay Area collection center and all eight data collection centers combined:

	San Francisco Bay Area center	All data collection centers
Breast Cancer Cases	865	4,742
Ovarian Cancer Cases	97	548
Endometrial Cancer Cases	100	672
Controls	820	4,754
Total Participants:	**1882**	**10,716**

Marin Residency Breakdown:
Not available in the reviewed published reports. The CASH study did not record county of residence at the time of data collection. They did inquire about the state and county where the respondent "spent most of [her] life" in under and over age 25 categories. County residence at time of diagnosis could be ascertained for cases by linking the CASH dataset to SEER data.

Data Collection Procedures:

Standardized one-hour interviews were administered to cases and controls in their homes.

Response Rates:

1985 SEER data reveal that, in general, case women who were not interviewed were diagnosed at later stages of disease and were less likely to survive than women who completed interviews (hence: if an exposure of interest is related to survival, results may have a selection bias). It is noted that non-response bias may be less of a concern for breast cancer cases because demographic differences among interviewed and non-interviewed cases were minimal (Wingo et al., 1988).

Bibliography of Associated Reports:

Wingo PA, Ory HW, Layde PM, Lee NC. The evaluation of the data collection process for a multicenter, population-based, case-control design. Am J Epidemiol 1998;128:206-217.

Dataset Name:

California Health Interview Survey (CHIS)

Dataset Source:

Data Access Center of the UCLA Center for Health Policy Research

Procedure for Obtaining Dataset:

State-level data are publicly available at the CHIS website (www.chis.ucla.edu). An application must be submitted to obtain individual-level data and data by geographic. Applications are reviewed by the CHIS Data Disclosure Review Committee, which makes recommendations to the CHIS Principal Investigator. The application may be sent to chis@ucla.edu.

Study Description/Abstract:

The 2001 California Health Interview Survey (CHIS) contains data on public health topics including health-related behaviors, health status and conditions, health insurance coverage, and access to health care services. CHIS is largely modeled after the National Health Interview Survey, and many questions were adapted from other national and state surveys as well as individual research projects that focused on population health. Data were collected for most California counties and for contiguous groupings of counties with small populations.

Years of Data Collection:

November, 2000 through September, 2001.

Study Design:

The CHIS interviewed one randomly selected adult in each of approximately 55,000 households. The sample was designed to provide population-based estimates for most California counties and all ethnic groups. It was also designed to optimize local-level estimates for counties with populations of 40,000 or more. Most households were contacted using random-digit-dialing techniques; some ethnic subgroups and geographic areas were oversampled.

Description of Participants:

Statistics by age and sex for the study and its Marin subgroup were not available through the reviewed resources. Statewide, 55,428 adults (age 18+) participated. 755 participants were in Marin County.

Data Collection Procedures:

Skilled interviewers called randomly selected households. Interviews were available in English, Spanish, Chinese (Cantonese and Mandarin), Vietnamese, Korean, or Khmer (Cambodian). Interviewers answered any questions about the interview and then randomly selected one adult in the house for the interview.

Response Rates:

Response rates were not available through the reviewed resources.

Bibliography of Associated Reports:

Dataset Name:

California Teachers Study (CTS)

Dataset Source:

Steering Committee of the California Teachers Study

Procedure for Obtaining Dataset:

This dataset is not a public use dataset.

Study Description/Abstract:

The California Teachers Study (CTS) is designed to document high breast cancer incidence rates of California teachers and to investigate emergent hypotheses in the etiology of breast and other cancers. It is a prospective study of 133,479 California female teachers and administrators. Established in 1995-1996, members of the California State Teachers Retirement System (CTRS) completed a detailed mailed questionnaire regarding possible risk factors for breast and other cancers. Cancer outcomes were identified by linkage with the California Cancer Registry.

Years of Data Collection:

1995-1996, 1997, 2000.

Study Design:

The CTS cohort is comprised of women who were active or retired teachers and were members of the CTRS in 1995. CTRS members are public school employees who teach or administrate at levels from kindergarten through community college. Where feasible, survey questions were drawn from established instruments.

Description of Participants:

Of the study's 133,479 participants, 15,660 were from the San Francisco Bay Area. Further subdivision by county and age was not available through resources reviewed.

Data Collection Procedures:

A mailed, self-administered questionnaire was sent to eligible women.

Response Rates:

Of 329,684 women contacted about the study, 133,479 participated. The response rate was 40.5%.

Bibliography of Associated Reports:

High breast cancer incidence rates among California teachers: results from the California Teachers Study. Cancer Causes and Control, 13:625-635, 2002.

Dataset Name:

California Women's Health Survey (CWHS)

Dataset Source:

The survey instrument, technical documentation, and datasets may be obtained from the Survey Research Group at srg@ccr.ca.gov.

Procedure for Obtaining Dataset:

Data requests are to be made of:
Dr. Bonnie Davis
Chief, SRG
Cancer Surveillance System
Ph: 916-779-0331
bonnie@ccr.ca.gov

Study Description/Abstract:

The survey consists of a core of demographic and health care access and health insurance coverage questions and is administered yearly. Questions cover topical areas specific to women's health, including breast and cervical cancer, mental health, family planning, breastfeeding, chronic illnesses, nutrition and obesity, sexual behavior, STDs, health information, caregiver responsibilities, and hormone replacement therapy.

Years of Data Collection:

Annually since 1997.

Study Design:

Sample design uses random-digit-dialed probability samples of the adult (aged 18+) female population. The annual statewide sample size is approximately 4,000 participants. The survey was weighted to approximate State of California population estimates. Participants are taken from a civilian, non-institutionalized adult population that resides in households with telephones.

Description of Participants:

Female participants by age and ethnicity:

Race/Ethnicity	18-24 y/o	25-34 y/o	35-44 y/o	45-54 y/o	55-64 y/o	65+ y/o	Total
White	609	1620	2257	2071	1340	2043	9940
Black	113	203	217	171	99	108	911
Hispanic	643	1438	1110	519	272	219	4201
Asian/Other	141	316	300	210	86	86	1139
Total	1506	3577	3884	2971	1797	2456	**16191**

Marin Residency Breakdown:

Marin residents comprise approximately 1% of respondents. Target annual sample size is 4,000.

2001 = 30 Marin residents
2000 = 37 Marin residents
1999 = 42 Marin residents
1998 = 40 Marin residents
1997 = 37 Marin residents
Total = 186 Marin residents / 16,191 Participants

Marin figures are based on what is available in the published reports only. Further refinement is possible only after obtaining the datasets.

Data Collection Procedures:

Telephone interviews:
Interviews were approximately 30 minutes in length and conducted in English and Spanish. Using the CATI system, interviewers read questions as they were displayed onscreen and keyed responses directly into the computer. Automatic data editing and coding programs were employed to increase the accuracy and speed of data entry.

Some questions in each interview may have been automatically skipped based on each participant's prior responses. Skip patterns are indicated on the survey instruments.

Response Rates:

	1997	1998	1999	2000	2001
Response Rate	67%	70%	81%	74%	[unreported]

These response rates were calculated using the following equation:

$$\frac{\text{Completed interviews x 100}}{\text{Refusals + Incompletes + Completes}} = \text{Response rate}$$

Bibliography of Associated Reports:

Dataset Name:

Family Registry for Breast Cancer (FRBC)

Dataset Source:

National Cancer Institute

The questionnaire may be obtained online:
http://www.cfr.epi.uci.edu/ic_registries/resources/Questionnaire%20pdfs/Breast/North%20California/FEMPROB_RV4.pdf

Procedure for Obtaining Dataset:

A brief but detailed proposal, including information about the study design and dataset request, is to be completed. Researchers must provide proof of approval for use of human subjects and pay appropriate expenses for section preparation and transfer of data.

Obtain the application form and related instructions from:
Mrs. Connie Galindo
Management Associate
Cancer Family Registries
Telephone: 301-594-3652
Fax: 301-435-5477
E-mail: galindoc@mail.nih.gov

Contact information from http://epi.grants.cancer.gov/BCFR/forms.html.

Also see http://epi.grants.cancer.gov/BCFR/Q&A.html#8.

Study Description/Abstract:

The purpose of the Family Registry for Breast Cancer (FRBC) is to organize a multi-center cooperative effort to collect pedigree information, epidemiological data and biological specimens from patients with a family history of breast cancer. Its goals are twofold: provide a collaborative infrastructure to facilitate interdisciplinary studies on the etiology of breast cancer and identify populations at all levels of risk for breast cancer that could benefit from new preventive strategies.

Years of Data Collection:

1996 - present

Study Design:

While some of the centers focus on the enrollment of high-risk families, the Northern California Cancer Center (NCCC) FRBC registry utilizes a broader approach. Families with one of the following criteria were invited to join: a male with breast cancer; a female with breast and/or ovarian cancer; or two or more first and second degree relatives diagnosed at any ages. Participants are asked to complete an epidemiology questionnaire, dietary history form, extensive family history, and donate biospecimens, including blood and urine. Pathology reports and tissue samples are obtained for breast and ovarian cancer cases. Pathology reports are requested to confirm the history of other cancers.

Description of Participants:

Specific information about the registrants, such as sex, age, race, and county residency details, requires investigator collaboration and submission of a dataset request (as described above). NCCC FRBC will eventually include data on about 1,250 Bay Area families.

Data Collection Procedures:

Families are recruited by telephone. Participants are asked to complete a questionnaire on family history of cancer occurrence, medical history, personal lifestyle factors and diet, and to provide a blood sample, and allow hospitals to provide access to previously collected pathologic specimens. All this information will be held anonymously in a central database for release to qualified researchers.

Response Rates:

Little information is available about response rates in the resources reviewed.

Bibliography of Associated Reports:

- Krieger N, Ashbury F, Cotterchio M, Macey J. A qualitative study of subject recruitment for familial cancer research. Annals Epidemiol, 2001;11:219-224.
- Daly MB, Offit K, Li F, Glendon G, Yaker A, West D et al. Participation in the Cooperative Family Registry for Breast Cancer Studies: issues of informed consent. J Natl Cancer Inst 2000;92:452-6.
- Di Prospero L, Seminsky M, Honeyford J, Doan B, Meschino W, Chart P, Warner E. Psychosocial issues following a positive genetic test for BRCA1 and BRCA2: Findings from a focus group and a needs assessment survey. Can Med Assoc J 2001 (In Press).
- DiSaia PJ, Brewster WR, Ziogas A, Anton-Culver H. Breast cancer survival and hormone replacement therapy: a cohort analysis. American Journal of Clinical Oncology (accepted for publication as of 5/31/00).
- Gilpin C, Carson N, Hunter AGW. A preliminary validation of a family history assessment form to select women at risk for breast or ovarian cancer for referral to a genetics center. Clin Genet 2000;58:299-308.
- Seminara D, ed. Innovative study designs and analytic approaches to the genetic epidemiology of cancer. J Natl Cancer Inst Monograph 1999;26:1-105.
- Senie R, Andrulis I, Daly M, Hopper J, Buys S, West D, Anton-Culver H. A unique resource for breast cancer research; The Cooperative Family Registry for Breast Cancer Studies. Eur J Hum Gene 2001; 9 Suppl 1:273.
- Senie R, Santella R, Ahsan H. The Metropolitan New York Registry and CFRBCS: Unique resources for breast cancer research. Ann Epidemiol 2000; 10:462.
- Spurdle, AB, Hopper, JL, Chen X, Dite GS, McCredie MRE, Giles GG, Venter DJ, Southey MC, Purdie DM and Chenevix-Trench G. The steroid 5 a-reductase type II repeat is not associated with risk of breast or ovarian cancer in Australian women. Submitted to Carcinogenesis.
- Sutherland HJ, Lacroix J, Knight J, Andrulis IL, Boyd NF and the Ontario Cancer Genetics Network. The Cooperative Familial Registry for Breast Cancer Studies: Design and first year recruitment rates in Ontario. J Clin Epidemiol 2001;54:93-98.

Dataset Name:

Marin County Health Survey (MCHS)

Dataset Source:

Contact:
Rochelle Ereman
Marin County Department of Health and Human Servces
415-499-3056

Procedure for Obtaining Dataset:

This dataset is not a public use dataset.

Study Description / Abstract:

The Marin County Health Survey (MCHS) is a cross-sectional study of health and health care among Marin County adults, children (age 0-18), and senior citizens (65 and older). Among its primary objectives are the assessment of health status, measurement of health-related behavior prevalences, and description of health care access and utilization among Marin County residents. Its design is intended to facilitate, where possible, comparison to state and national populations.

Years of Data Collection:

June, 2001 – October, 2001

Study Design:

The MCHS was a cross-sectional study of Marin County residents. It employed an unrestricted random digit dial sampling methodology of county residences for its Main Adult Survey; the main adult in households with more than one eligible participant (ages 18+) was systematically selected using a "most recent birthday" technique. The Child Follow-on Survey was conducted with the mother or primary caregiver for those households that reported one or more children in the residence under the age of 18. Seniors age 65 or older who were interviewed in the Main Adult Survey were asked to participate in the Senior Follow-on Survey.

Description of Participants:

All Participants were residents of Marin County at the time of recruitment.

Female participants by age:

	18-24 y/o	25-34 y/o	35-44 y/o	45-54 y/o	55-64 y/o	65-74 y/o	75-84 y/o	85+ y/o	Total
No. Participants	127	387	659	683	486	247	195	38	**2,822**

Data Collection Procedures:

Computer-assisted telephone interviews:
The MCHS was administered by computer-assisted telephone interviewing (CATI). Field Research Corporation contractors conducted interviews in English (4,730—98%) and Spanish (91—2%). The Main Adult instrument was

designed to be no more than 30 minutes long, with the Child and Senior Follow-on instruments no more than 20 minutes long.

Response Rates:

The MCHS Sample Disposition Report indicates that 13,818 eligible numbers were found among 31,692 numbers identified by RDD. Among the eligible numbers, 3,194 could not be completed during the field period, 5,803 refused, and 4,821 completed the Main Adult survey. Child Follow-on interviews were conducted with 898 Participants (87% of main adults who had eligible children). Senior Follow-on interviews were conducted with 700 seniors (88% of households in which the main adult was 65 or older).

Bibliography of Associated Reports:

Appendix A

TRIFS Dataset Summary Table

	ARFS		CTS			BRFS	CWHS	Buck				FRBC	CASH	CHIS	MCHS
	1	2	B	1	2			1	2	S	3				
Sex	✓	✓	✓	✓	✓	84-02	✓	✓	✓	✓	✓	✓	✓	✓	✓
Age	✓	✓	✓			84-02	✓	✓	✓	✓	✓	✓	✓	✓	✓
Age at BC Diagnosis	✓		✓									✓			
Country of Origin	✓		✓				✓					✓	✓	✓	✓
History [1] of Cancer in One Breast	✓		✓				✓	✓	✓	✓	✓	✓		✓	✓
Family History	✓		✓										✓	✓	✓
Biopsy	✓		✓									✓	✓	✓	
Radiation Exposure	✓			✓								✓			✓
Screening (e.g, BSE, mammography, thermography)	✓			✓		87-02	✓			✓		✓	✓	✓	✓
History of Ovarian Cancer	✓		✓		✓			✓	✓	✓	✓	✓	✓	✓	✓
History of Endometrial Cancer	✓		✓		✓			✓	✓	✓	✓	✓	✓	✓	✓
Benign Breast Disease	✓		✓					✓	✓	✓	✓	✓		✓	
Ovary Removal	✓		✓		✓					✓		✓	✓	✓	
Parity	✓		✓	✓			✓					✓	✓	✓	✓
Age at First Full-term Pregnancy	✓		✓				✓					✓	✓	✓	✓
Age at Menarche	✓		✓									✓	✓	✓	
Age at Menopause	✓		✓		✓					✓		✓	✓		
Oral Contraceptive Use	✓		✓		✓	98-00, 02	✓					✓	✓		✓

ARFS: Adolescent Risk Factor Study
CTS: California Teachers Study- Baseline, I, II
BRFS: Behavioral Risk Factor Survey. Figures indicate years of administration during which related questions were present.
CWHS: California Women's Health Study
Buck: Buck Center Study- "Health and Functioning in Marin County". Baseline, Exam II, Surveillance, Exam III
FRBC: Family Registry for Breast Cancer
CASH: Cancer and Steroid Hormone Study
CHIS: California Health Interview Survey
MCHS: Marin Community Health Survey

(last updated 11/14/02)

LBNL-51905

Appendix A

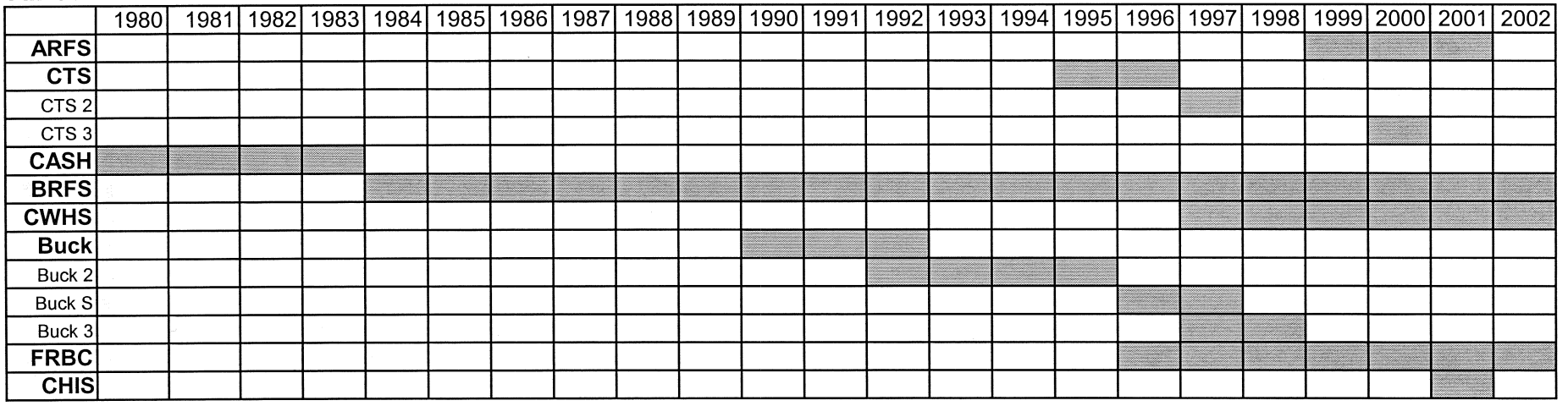
TRIFS Dataset Summary Table

	ARFS		CTS			BRFS	CWHS	Buck				FRBC	CASH	CHIS	MCHS
	1	2	B	1	2			1	2	S	3				
Hormones (other than oral contraceptives)	✓		✓		✓	96-00, 02	✓	✓	✓	✓	✓	✓	✓	✓	✓
SES	✓					84-02	✓	✓						✓	✓
Education	✓					84-02	✓	✓				✓	✓	✓	✓
Income						84-02	✓	✓			✓		✓	✓	✓
Marital Status	✓		✓		✓	84-02	✓	✓	✓		✓	✓	✓	✓	✓
Urban/Rural residence	✓	✓													
Geographic location (e.g., region, Lat./Long.)	✓	✓	✓	✓	✓	88-02	✓							✓	✓
Race/Ethnicity	✓		✓			84-02	✓	✓				✓	✓	✓	✓
Religion	✓										✓	✓	✓		
Obesity (BMI)	✓		✓	✓		84-02	✓	✓	✓			✓	✓		✓
Height	✓		✓	✓		84-02	✓	✓	✓			✓	✓		✓
Weight	✓		✓	✓		84-02	✓	✓	✓			✓	✓		✓
Alcohol	✓		✓			84-02	✓	✓				✓	✓	✓	✓
Smoking	✓		✓	✓		84-02	✓	✓	✓		✓	✓	✓	✓	✓
Diet	✓		✓	✓		84-02			✓	✓				✓	✓
Exercise	✓		✓			84-02		✓	✓		✓	✓		✓	✓
Occupation	✓		✓		✓		✓	✓	✓					✓	✓
NAF															
Breastfeeding	✓											✓			✓
Mammographic Pattern															
Environmental Factors (Yes/No)	✓		✓												

(last updated 11/14/02)

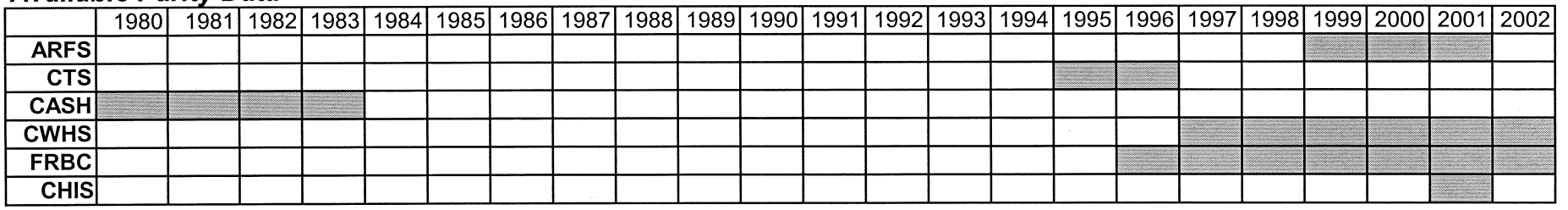
TRiFS Datasets: Timelines of Study Collection Efforts

All Studies

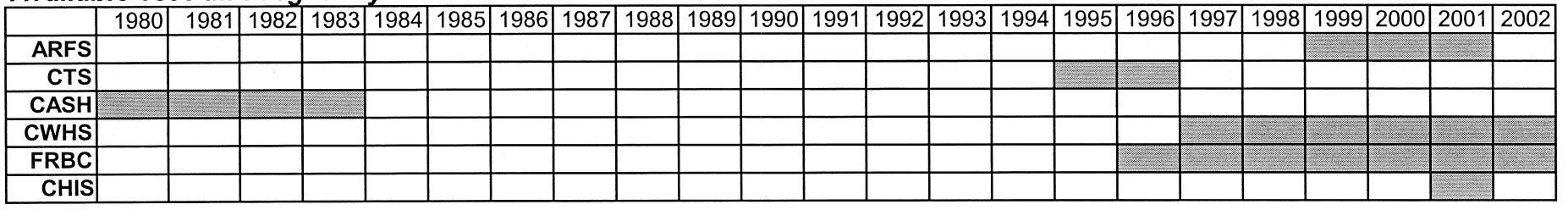


TRiFS Datasets: Timelines of Study Collection Efforts by Risk Factor

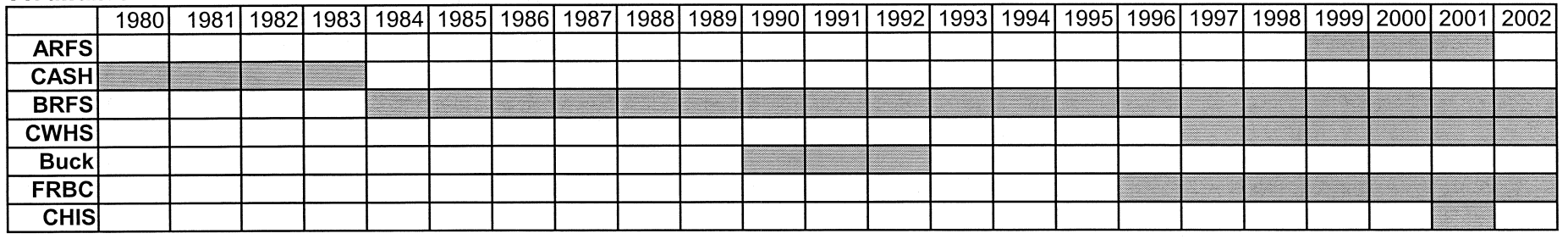
Available Parity Data



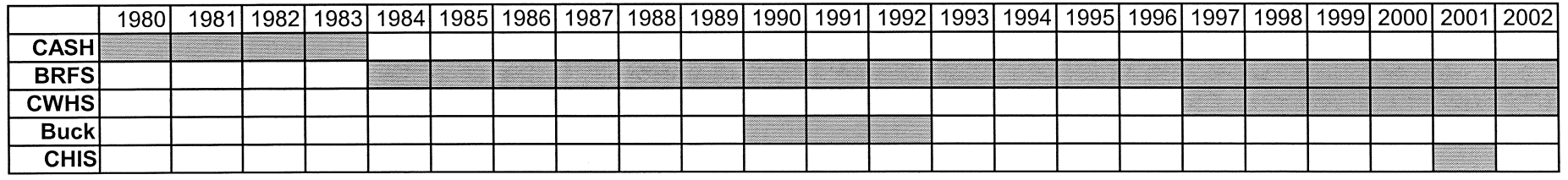
Available 1st-Full-Pregnancy Data



Available Education Data



Available Income Data



APPENDIX B - Slides from presentation given at Marin County Town Meeting on 5/29/03

Traditional Risk Factor Study

Christine A. Erdmann, Ph.D., MPH
Exposure and Risk Analysis Group
Lawrence Berkeley National Laboratory

William A. Satariano, Ph.D.
Division of Public Health Biology & Epidemiology
School of Public Health
University of California at Berkeley

Ying Qing Chen, Ph.D.
Division of Biostatistics
School of Public Health
University of California at Berkeley

LAWRENCE BERKELEY NATIONAL LABORATORY

Traditional Risk Factor Study (TRIFS)

Research Question: What proportion of breast cancer cases in Marin County is attributable to traditional breast cancer risk factors?

Objectives:

1. Describe female breast cancer risk factor distributions in Marin County using previously collected individual-level data.
2. Compare breast cancer risk factor distributions for Marin County with those of other California Counties and the State.

Previous Studies

Robbins et al. (1997) → Limitations:
- Regional differences accounted for by regional differences in age, parity, age at 1st birth, breastfeeding, age at menarche, age at menopause, and alcohol use

Limitations:
- Restricted age group, 22-55 years
- Marin grouped with other SFBA counties

Prehn & West (1998) → Limitations:
- Rates in Marin comparable to those in areas with similar characteristics (age, parity, income, education)

Limitations:
- Group level data
- Proxy measures of risk factors

TRIFS Tasks

- Reviewed existing datasets with individual-level risk factor data for Marin County and other CA counties
- Selected 2 datasets for detailed study:
  1. Adolescent Risk Factor Study (ARFS)
  2. California Health Interview Survey (CHIS)
- Calculated prevalence & population attributable fraction estimates for traditional breast cancer risk factors using these datasets

Traditional Breast Cancer Risk Factors

Variable	Higher Risk Group Definition
Age at 1st birth	Age at 1st birth ≥ 30 years, includes nulliparous
Family history	Breast cancer in one or more 1st degree relatives
Parity	Nulliparous
Age at menarche	Age at menarche < 12 years
Body mass index	Body mass index > 27 among postmenopausal women
Age at menopause	Age at menopause ≥ 55 years
Education	Highest educational level at least a Bachelor's degree
Socioeconomic status	Higher socioeconomic status

Population Attributable Fraction (PAF)

Rockhill et al. AJPH (1998)

Definition:
The proportion of cases that would be prevented if the risk factor were eliminated.

Formula: $\frac{(P_e)(RR - 1)}{(P_e)(RR - 1) + 1}$

where P_e = *proportion of women with the risk factor*

RR = relative risk estimate for risk factor

$= \frac{\text{proportion of cases with factor}}{\text{proportion of cases without factor}}$

= strength of association between the risk factor & the disease

Appendix B

Example Calculation

Imaginary Risk Factor:
Handled 2 or more frogs during adolescence

Pe = 0.30, RR = 1.6

$$\text{Population Attributable Fraction (PAF)} = \frac{(P_e)(RR - 1)}{(P_e)(RR - 1) + 1} = \frac{(0.30)(1.6 - 1)}{(0.30)(1.6 - 1) + 1} = 0.15$$

Example Interpretation

Population Attributable Fraction (PAF) = 0.15

Correct Interpretation:

If all females in Marin County never touched a frog during adolescence, then 15% of breast cancer cases in the female population of Marin County would be avoided.

Prevalences & Population Attributable Fractions for Marin County

Risk Factor	Pe	RR	PAF x 100
Age at 1st birth ≥30 years	0.51	1.9	31.3
Family history	0.20	2.6	24.2
Parity - nulliparous	0.27	1.9	19.3
Age at menarche <12 years	0.22	1.2	4.2
Body mass index (postmenopause) >27	0.40	1.1-1.4	3.9-12.9
Age at menopause ≥ 55 years	0.14	1.1-2.0	1.4-12.0
Education ≥ B.A./B.S.	0.60	1.4-2.3	19.3-43.7
Higher socioeconomic status (SES)	0.53	1.1-2.0	5.0-34.6

Pe = prevalence of risk factor among Marin County females

RR = relative risk estimates from published sources

PAF x 100 = population attributable fraction expressed as a percentage

Prevalences & Population Attributable Fractions for Marin County

Risk Factor	Pe	RR	PAF x 100
Age at 1st birth ≥30 years	0.51	1.9	31.3
Family history	0.20	2.6	24.2
Parity - nulliparous	0.27	1.9	19.3
Age at menarche <12 years	0.22	1.2	4.2
Body mass index (postmenopause) >27	0.40	1.1-1.4	3.9-12.9
Age at menopause ≥ 55 years	0.14	1.1-2.0	1.4-12.0
Education ≥ B.A./B.S.	0.60	1.4-2.3	19.3-43.7
Higher socioeconomic status (SES)	0.53	1.1-2.0	5.0-34.6

Correct Interpretation: If all postmenopausal females in Marin County had a body mass index < 21, then approximately 3.9-12.9% of breast cancer cases in the female population of Marin County would be avoided.

Population Attributable Fraction for Marin County

Risk Factor Combination	PAF x 100
Age at 1st birth ≥ 30 years Family history Age at menarche <12 years Age at menopause ≥55 years Body mass index (postmenopause) >27	51.3-57.6

Correct Interpretation: If these breast cancer risk factors were eliminated in Marin County, then about 50% of breast cancer cases in Marin County would be avoided.

Warnings

- Because many risk factors are correlated, population attributable fraction (PAF) estimates calculated for single risk factors cannot be summed
- Selection of cutpoints for prevalence estimates can greatly impact the result
- PAFs for factors that are not easily modifiable are less useful measures for prevention
- Actual reduction in disease burden after removal of the risk factor assumes that the risk factor is causally related to the disease
- PAF is easily misinterpreted

Population Attributable Fraction

Correct Interpretation:

If certain traditional breast cancer risk factors were eliminated in Marin County, then about 50% of breast cancer cases would be avoided in Marin County.

Incorrect Interpretation:

~~Only 50% percent of Marin County women have any known breast cancer risk factors.~~

Approximately 84% of Marin women women have at least one of the following risk factors: earlier age at menarche, later age at 1st birth, family history of BC, later age at menopause, higher postmenopausal BMI.

What about the remaining 50% of breast cancer cases in Marin County?

- Hormone replacement therapy (HRT)?
- Alcohol?
- Other environmental factors?
- Wrensch et al. (2003) found no notable differences between cases and controls for
  - –age 1st lived in Marin County,
  - –lifetime years in Marin County, or
  - –lifetime years in the San Francisco Bay Area
- Same unknown factors that account for unexplained 50% in other populations?

Summary

- Considered in combination, traditional breast cancer risk factors appear to account for about 50% of breast cancer cases in Marin County
- Later age at 1st birth and nulliparity after age 30 appear to account for about 1/3 of breast cancer cases in Marin County
- 84% of Marin County women are exposed to at least one traditional breast cancer risk factor

Acknowledgments

People:
Margaret Wrensch, Ph.D., UCSF
Terri Chew, MPH, UCSF
Lee Habte, CHIS, UCLA
Colin Leary
Mike Sohn, Ph.D.
Matt McLeod, Ph.D.
Marion Russell
Agnes Lobscheid
} Lawrence Berkeley National Laboratory

Funding:

This work was made possible by funds received from Agreement No. 00-91462 with the California Department of health Services, Cancer Research Section. These funds were administered through the Marin County Department of Health and Human Services.

Appendix B